IEEE Intl. Symp. Intell. Control (Workshop on Architectures for Semiotic Modeling and Situation Analysis in Large Complex Systems, Orgs.: J. Albus, A Meystel, D Pospelov, T. Reader), Monterey, CA. August, 1995.

Human-Machine Interaction in Telerobotic Systems and Issues in the Architecture of Intelligent and Cooperative Control

Paul S. Schenker
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive/ MS 198-219
Pasadena. CA 91109
EMAIL: schenker@telerobotics.jpl.nasa.gov

Gerard "I". McKee
[University of Reading
Department of Computer Science
Whiteknights
Reading RG62AY, UK
I.MAIL.: Gerard.McKee@reading.ac.uk

Abstract

The design of architectures for robust, intelligent telerobotic interactions is challenging, both on representational and algorithmic grounds. From a structural standpoint, a telerobotic system is inherently *distributed. The operator and robot are separately situated, one in a physical world and the other in partially modeled manifestation. The communication of task state information and control is contaminated by* systematic and random errors, and maybe *time-delayed. From a representational standpoint, multiple task resolutions are required, and uncertainty is inherent. Different abstractions of the task domain. operator actions, their signification, and* domain constraint consistency *mus(be* maintained. *Both continuous state and discrete-linguistic task models must be accommodated and coordinated. Finally, telerobotic systems operate in non-stationary, highly diverse physical environment. We overview these* issues in the *context of this workshop on intelligent controls, linguistic structures and their descriptive and analytic roles in future synthesis of large complex systems. We illustrate some of the issues with design* examples drawn from recent *work on intelligent and cooperative control of remote multi-camera viewing and dexterous manipulation.*

Keywords: robotics, teleoperation, telerobotics, human-machine interface, intelligent control, graphics user interface, virtual reality, sensor planning

1 Introduction& Background

Telerobots, such as the example pictured in **Figure 1, are** systems in which a human agent controls a remote robotic one. The motivations for developing telerobotic systems are several-fold and include extending human expertise to distant frontiers (space/underseas/telesurgeries), distancing human expertise from dangerous environments (nuclear/military), and enhancing human expertise through machine assistance (micro-assembly/biomedicine). In most situations where a telerobot would be used, the task is comparable to one that would be performed by human means. It follows that telerobots are among the most complex systems considered to date. -- complex as engineering artifact, as computational construct, and as man-machine interface.

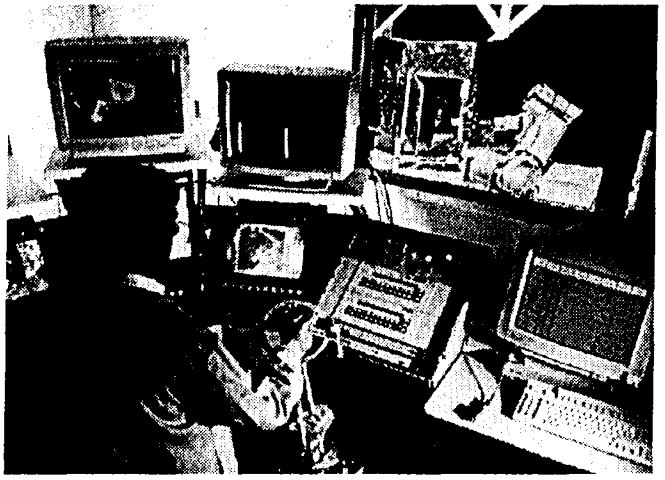

Fig. 1: Operator conducts simulated telerobotic servicing of a satellite (Jet Propulsion Laboratory)

In this section wc outline human-machine interaction in telerobotic systems, emphasizing the scope and hierarchy of task information & control structure shared by human and robot agents. Graphically mediated interaction is one important way of signifying operator intent and task status. We discuss this idea in *Section 2*, presenting several experimental examples. In *Section.? we* examine the use of mixed graphic and linguistic constructs, illustrating this approach with a recently developed telerobotic architecture for intelligent viewing control. Finally, in *Section* 4, we conclude with comments on open issues and future research ties between telerobotic architectures, linguistic structures, controls, and **the** coherent synthesis of large complex systems.

From the viewpoint of operational abstraction, telerobotic control modes span teleoperation to supervisory automation. In *teleoperation*, the operator signifies by manual input a desired set of motion trajectories in the remote [ask space. A joy-stick like device (hand-controller) enables independent motion of each robot joint ("joint space control") or task space positioning of the robot arm tip ("Cartesian control"). The operator monitors robot motion by remote camera views. visually servoing his hand input motions in response 10 signified discrepancies from the envisioned ideal trajectory. When the robot is brought into contact with the remote task work pieces, it is desirable to give indication of exerted forces. This can be done by instrumenting the robot hand/wrist with strain gauges and presenting a properly transformed force to the operator's hand controller, via a motor drive. Such force *telepresence* prevents damage to robot or work piece, and more fundamentally, facilitates manual dexterity. E.g., teleoperative task times with kinesthetic force-feedback typically are 40-60% of those without. There are many engineering and manual controls design challenges in such high dexterity kinesthetic force-reflecting *systems* [13, *50, 56]*, and this remains an active research area. Two key design issues, further discussed in Sections 2 *and.? are: 1)* effects of intervening communication time-delay, 2) adequacy of remote visual task presentation to efficient telerobotic operation.

At the other operative extreme, *supervisory automated control, the* operator/robot discourse is discretized and inherently symbolic. Activities of this type are distinguished from robot ic automation primarily by rest-time operator/robot interaction and the degree to which human and robot agents depend *on* simultaneous, consistent model and state knowledge for task prediction, execution, and verification. "Shared controls" illustrate both the complexity and advantages of this regime. In such implementations, a task function is decomposed into partial degrees of freedom (perceptual or manipulative), some of which are automated and some teleoperative. There are different ways in which the operator can cognitively and physically interact with the robot to communicate intent, and a hierarchy of physical, sensorial, iconic, and linguistic signifiers by which a hi-lateral distributed task representation is established. Albus et al. [7-9], Conway et at. [19], Paul et al. [44], and Schenker et al. [34, 51] outline a few of the telerobotic architectures that have been used in this context, and references [1] and [56] provide related perspectives on robotic automation and supervisory control.

2 Graphics-Based Telerobotic Interaction

There is a recent trend to use of graphic and iconic environments as a primary means of human-machine interaction in telerobot ic systems [10, 18, 22, 31, 33, 35, 63], Such user interfaces play *simulative* roles as well as *executive* ones [29, 30, 35, 42]. They are conceptually appealing, encourage a functional orientation to system programming, and enable high efficiency information transfer. Progress in 3-D real time graphics/VR technologies has accelerated this developmental trend. Schenker (52] has summarized an informal taxonomy of graphics-based operator interactions, characterizing these as *manually servoed, manually guided, and manually designated*. We next summarize these operational modes and illustrate them with experimental examples.

2.1 Manually Servoed operations

Definition: The operator's input is spatially and temporally continuous, and the 3-D graphic display presentation is real-time and instantaneous; the operator's analog positioning control inputs arc either simultaneously or subsequently issued to the robot itself. The simulated task presentation to the operator may include other sensory modes than visual, such as modeled force feedback [29]. *Scope: "Preview"* and "predictive" graphic displays are examples, per references [30-32, 36,57]. Those remote robotics applications wherein the operator's motion input to the graphics interface and distant robot controller are issued simultaneously (versus a "record-preview-playback mode") usuall y assume a synchronous communication channel. This approach is potentially effective for time-delayed teleoperation over same in well-modeled, but casually structured tasking environments.

Example: As one example of manually-servoed operations we sketch a recent development supporting time-delayed ground-to-orbit space telerobotic servicing. This work both illustrates some of the conceptual problems introduccd by communication time-delay, as well as the task representation roles played by 3-D graphics. Note that efficient, reliable teleoperati ve tasking depends critically on the operator's eye-to-hand motion coordination -- the synchronization of an operator's causal hand controller input and visual display of resulting effect on robot and task, Such coordination significantly degrades at fractions of a second perceptual time delay, and is essentially lost beyond a second, forcing the adoption of highly ineffectual move-and-wait tasking strategies [57]. This problem can be ameliorated by restoring the operat or's 3-D instantaneous perception with a graphics-based task model, which must be properly calibrated to both the real task geometries and operator's viewing perspective. Sheridan and students at MIT [56, 57] motivated early work of this type by introducing a stick-figure robot "predictor" overlay graphic on the operator's camera viewing screen; the graphics were only qualitatively correlated with the remote view, and modeled just the robot. More recently, JPL investigators have developed manually-servoed predictive graphics interfaces [29-32] which are rigorously calibrated to the remote task. The essence of these developments is a 3-D real-time graphics interface in which high-fidelity real-time renderings in either wire-frame or solid modeled (shaded polygonal) surfaces can bc shown in a transparent graphics-on-video overlay mode, and in which salient geometric features of both the robot and workspace objects are modelled. The graph ics can be rendered at flicker-free NTSC monocular or stereo display rates, in full geometric calibration to both the actual multi-camera views and object positions-and-poses. Once calibration is established, it is also possible, within limits of modeling fidelity, to *synthesize realistic task view presentations from an arbitrary viewpoint.* This is a very important capability 10 have when performing robotic operations in obstructive. limited access areas -- e.g., for inspection, servicing, and salvaging operations in-and-about complex platform structures [4].

Jet Propulsion Laboratory (JPL, Pasadena, CA), working with the Goddard Space Flight Center (GSFC, Greenbelt, MD) has used such a preview/predictive graphics interface to simulate space satellite servicing under time delay [30]. In [his experimentation, a JPL-based operator controlled the changeout of a satellite-mounted Orbital Replacement Unit (ORU) at GSFC. The GSFC remote robotic workcell was composed of single 7-d.o.f. arm with an attached 18" power latch-driver Lightweight Servicing Tool (LST), and operated under Cartesian space position control in 6-d.o.f. with a generalized "wrench" position compliance referred to a wrist force-torque sensor. The remote site controls also incorporated a novel GSFC-developed "capaciflector" proximity sensor which can be used for ORU fine positioning alignments, and this was tested during some aspects of operation. Four task views were available from the remote cameras: overhead-front wide and zoom, right-oblique, and right-side task presentations. Communications between the JPL and GSFC sites utilized the NASA Select NTSC satellite communications (30 fps) channel for video, and TCP/IP bidirectional socket/Internet data links for command & control. The aggregate JPL-GSFC round trip time delays varied between 4-to- 15 seconds. The JPL operator's primary interaction with the remote task during execution was the visually calibrated preview display and his manual inputs to the 6-d.o.f. robot hand-controller (see earlier Figure 1). **Figure** 2 shows a representative screen preview display during an actual task

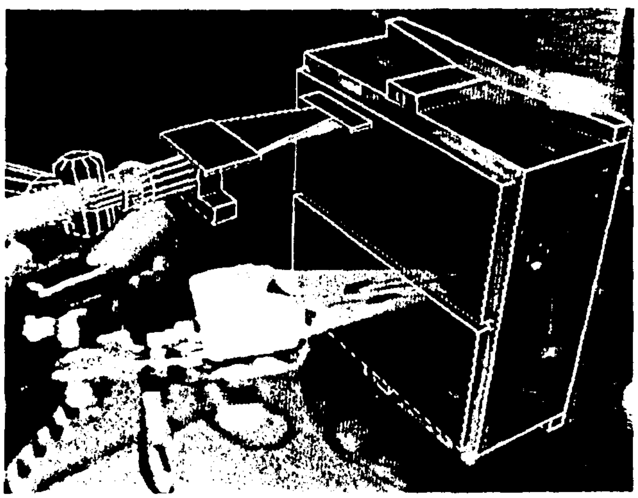

Fig. 2: Graphics Preview/Predictive Display

execution -- the operator has generated and previewed a hypothesized trajectory that will bring the ORU servicing tool into alignment with its insertion axis. and remote site robot motion to the predicted pose is ready to commence.

The non-real time functions of task calibration, robot control configuration, execution sequencing, and general command status reporting are performed at a Silicon Graphics workstation positioned to the operator's rear. The top-level workstation screen layout is shown in **Figure** 3, below.

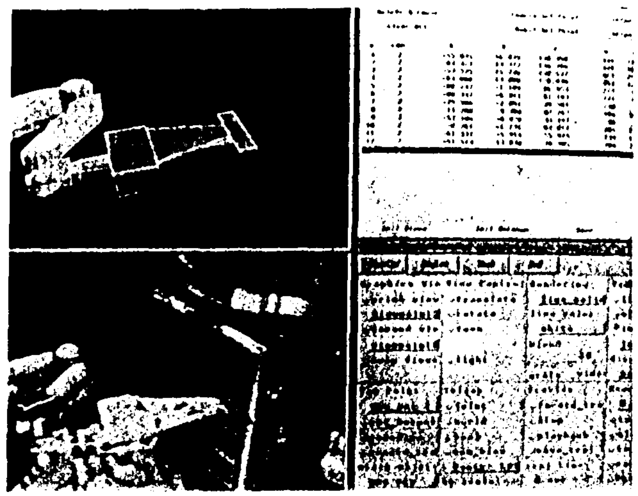

Fig. 3: Graphics Preview/Predictive Control Interface (Viewing Calibration in Progress)

Note the essence of this overall approach to telerobotic tasking is *analogical* -- motion control, command communication, and perceptual representation of the task at hand are based on continuous state information and model constructs.

2.2 Manually Guided Operations:

Definition: the operator's input is spatially and temporally continuous, and the 3-D graphic display presentation is again real-time and instantaneous (possibly including simulated forces or synthesized sensor feedback from the remote site). However, positioning control inputs to the remote robot are now the result of a higher level, computer-mediated symbolic reasoning about of the operator's basic free space and force contact motions, possibly in combination with robot site sensor-derived logical conditions or behaviors. For example, given a sufficiently deep task model, an operator-site symbolic interpreter can be used to generate a low-level autonomous command sequence (of motion primitives) that is issued asynchronously relative to the operator's input, and syntactically parsed back to continuous state analog motion-and-force controls at the remote site. *Scope:* "Teleprogramming" interfaces [24, 44] fal I into this category, as do also some feature-based learning and associative, sensor-referenced control schemes [26-28, 63] conditioned by a human teacher. By definition, this approach to telerobotic operations is more tolerant of lime-varying communication channels, and potentially effective for "teleoperator-like" tasking over same in well-modeled. reasonably well-slmcturcd environments.

Example: R. Paul and students at The University of Pennsylvania have pioneered a number of "teleprogramming" developments [24. 49, 60] enabling a hybridized system control loop and information exchange between human and robot agents under time delay. **As a logical extension** of continuous state predictive displays and task simulators, these developments are salient for several reasons. First, as noted above. this human-machine interaction paradigm directly addresses the problem of robust tasking over communication channels that are time-delayed, asynchronous, and implicitly degraded -- by "symbolically" breaking the continuous signal loop between operator and robot into two loosely coupled local and remote feedback control loops. Second. it retains a perceptually transparent analog task representation for both forward and backward operator interactions, while also providing a basis for discrete state iconic presentation/verification when desired [61]. Third, it explicitly inquires into the problem of chunking telerobotic perceptual and control functions into a higher level syntactic abstractions and semantics, e.g., as associated with the various sensor-referenced primitives for free-space, guarded, and full contact motions. Stein's recent thesis [61] has generalized the teleprogramming concept to incorporate behavior-based adaptation [14, 15, 23] of the robot agent's control in response to real-time sensor data. Working in collaboration, UPenn researchers and JPL recent] y implemented and demonstrated at JPL such a layered robot control architecture, integrating it with the real-time robot controller and remote command interfaces of the robot workcell shown in Figures 1 and 8. When active, this behavioral control replaces more conventional continuous state hybrid position/force and compliance controls [1, 20, 21] as often used to correct *quantitative* variations in robot force and position along various axes of robot tool or gripper contact with an object, and can autonomously compensate and correct for undesirable *qualitative* changes in the task state, e.g., as determined by the robot force sensors. For example, the controller can assist a time-delayed operator in cleating with sudden, unpredictable disturbances and variations in contact with a workpiece being serviced, or object encountered. In 1994, an operator at UPenn (Philadelphia, PA), as depicted in **Figure 4,** successfully controlled the JPL robot under time delay to puncture and slice a Kapton tape seam securing satellite thermal blankets about a replica ORU access panel door [62]. Guided by the UPenn operator's analog positioning command inputs relative to his graphically model of the remote JPL workspace, the behavior-based controller robustly managed multiple, unpredictable metal-to-metal sidewall contacts as a cut ting tool traveled laterall y along a 2 mm. wide groove of a continuous linear 40 cm. path sweep.

Fig. 4: *Operator at UPenn Commands JPL Robot*

Figure 5 shows the experimental setting for this task. Communication time delays were intermittent bet ween 3-to-15 seconds, with an average latency of 6 seconds. Note that such tasks are challenging even for manual teleoperation without time delay [20, 21].

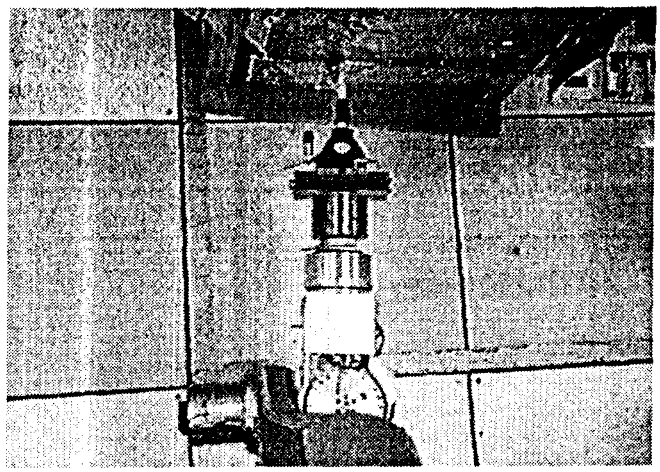

Fig. 5: *Satellite Servicing under Teleprogrammed Control*

2.3 Manually Designated Operations

Definition: the operator's inputs are spatially designated and temporally ordered symbolic actions, such as 3-D *point-to, select, drag, drop,* and other primitives with conceptual counterparts in 2-D GUIs [55]. The operational essence of this approach is a *visual programming of discrete task events,* and the resulting control strategy is dictated both by the semantics of the 3-D iconic primitives, and underlying supervisory control structure [1 1] available to support them.

In a broader definition, such iconically driven multi-resolution approaches *[22. 33. 51. 65]* might hybridize and functionally represent features of both supervisory and behavioral sensor-referenced controls, perhaps using the former at multiple levels to resolve-and-index emergent conflicts-and-skills of the latter; for example, sce discussions and references of [591. *Scope:* "Graphical programming" interfaces [35] typify this approach, which is limited at present by task representation. machine perception, spatial planning, dexterous controls, etc. -- a key issue *being* the degree to which these underlying techniques/tools generalize. The approach is potentially most effective in domain specific, well modeled, highly structured tasking scenarios, wherein the operator can flexibly and transparently intervene to reparameterize and reinitialize a sequence when necessary, as well as invoke strong prior task knowledge.

Examples: There area number of examples for this class of man-machine interaction, the scope of which is highly dependent on the domain definition and physical abstraction of task complexity. Indeed a number of CAD modeling, robot assembly programming, and image processing systems routinely use similar iconic front ends to facilitate operator interaction with their underlying linguistic constructs. We note in passing two rather unique developments. One is the ***MEISTER*** *(Model* Enhanced Intelligent and Skillful TEleoperational Robot) system [26, 48] of Electrotechnical Laboratory, in which the emphasis has been higher-level intelligent and cooperative human-machine control of telerobotic tasks. In one such demonstrated task -- robotic chemical assay by flame test -- a robot under supervised autonomy sets-ups. pulverizes, samples and flame-tests chemical substances. with the operator intervening to graphically re-designate locations of desired actions or teleoperate to deal with task anomalies. ETL has explicitly developed the MEISTER system architecture to enable ***multiple*** *level resolution* interaction. A key enabling design feature of this architecture is the embedding of environmental and control knowledge within a collect ion oft ask-oriented object models, wherein the model representation itself is fundamentally "object-orientCd." Viz., each object model contains self-knowledge such as position and orientation with respect to world coordinates ("object localization") and its affixment relationships to other objects. The object models embed both generic and specific handling knowledge, such that the commanding of a control operation, e.g., pick_and_place, invokes a linked hierarchy of processes, including the automatic sequencing of basic camera viewing primitives. MEISTER incorporates a motion understanding system that interprets and maintains consistent world model representations in response to multi-level human interventions into motion control. The second example is development by ATR Communication Laboratories of a system for object manipulations and layout in 3-D virtual workspaces [41]. Conceived as environment for cooperative workgroup design by teleconferencing, this system uses the combination, and probabilistic intersection (by means of spatial distributions on positional indicator words like **here, above, left,** UC.), of abstract natural language and instrumented hand pointing gestures to recognize the signified purpose of an agent. The system transforms verbal semantics into a spatial regions, and utilizes object specific knowledge (attachment relationships, holonomic and non-holonomic motion constraint axes, UC.) to resolve unambiguous object motions and transformations/affi xments -- e.g.. **move_the_**desk_ **to_the_front_left_of_the_table.**

3 Telerobotics & Intelligent Viewing Control

We next briefly overview a knowledge-based multi-resolution architecture for computer assisted sequencing of multiple cameras during telerobotic tasks, as well as graphics based synthesis/prediction of arbitrary task views -- two problems related to *sensor planning* in computer vision [6, 64]. The purpose of this section, along with its intrinsic value as a new telerobotic system design description, is to concretely illustrate some further issues of higher-level human-machine interaction in telerobots.

"Intelligent Viewing Control **(IVC),"** as we refer to this JPL-based development 153], utilizes mixed graphical and linguistic representations, and is one instantiation of a broader set of system concepts for multiple agent viewing control and cooperation that McKee and Schenker have reported elsewhere [38, 39], including a companion paper of this workshop [40]. The motivation of this work, beyond the study of intelligent system architectures, is quite practical: The cognitive workload of managing and integrating remote viewing resources -- the positioning, sequencing, panning, zooming, and focussing of several cameras -- is sufficiently high that a second operator is often required solely for this function [17, 20]. This additional workload and personnel assignment is often unacceptable in resource-limited mission environments like undersea and space telerobotic operations. Even when the use of a dedicated camera operator is practical, the robot operator still must maintain a complex verbal dialogue with camera coordination activities, imposing a secondary task load that distracts him from the primary manipulation objective. ~'bus, technical developments which automate 3-D task viewing procedures and increase their operational productivity are functionally significant [25, 361. Further, developments that improve 3-D task comprehension under adverse viewing conditions **(e.g..** providing visual/virtual aides to compensate limited or obstructed views, low acuity, etc.) are important toward improved tasking operational flexibility and safety [4].

3.1 **IVC** Functional **Architecture**

Definition: per **Figure 6**, overleaf, IVC knowledge (task geometry, task sequence, sensor and control models, etc.) is at any given time distributed among various logical *devices.* The IVC hub provides direct access to these system devices and top-level access to all distributed *knowledge objects.*

Within Figure 6, "ETL" refers to control/communication interfaces with Electrotechnical Laboratory. Tsukuba, Japan. with whom JPL is conducting trans-Pacific experimentation on telerobotic servicing and assembly tasks [54, 68].

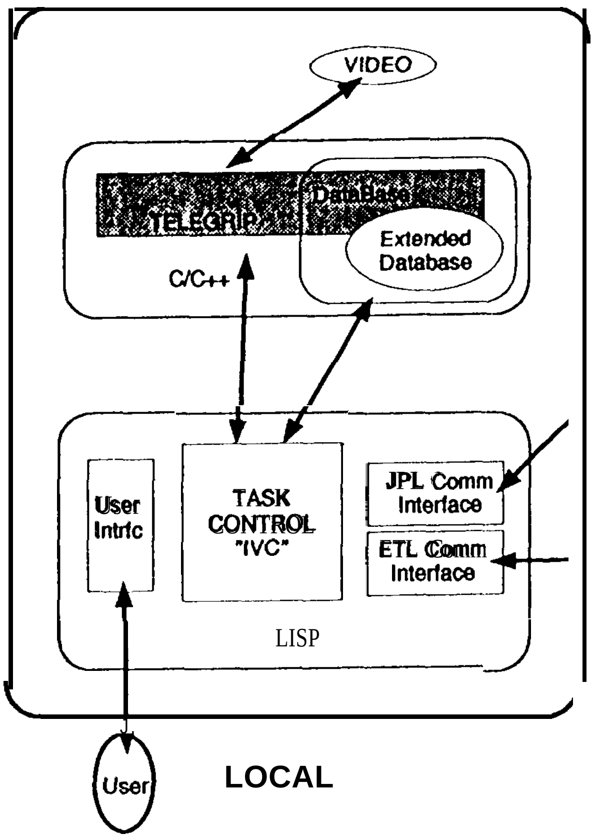

Fig 6: Block Architecture of IVC Sub-System

The user interface provides graphic and textual entry for *low-level* commanding of all system devices, *action-level* commanding of individual mechanisms, and *semantic-level* commanding of coordinated actions involving multiple mechanisms or multiple task steps (Terminology in *italics* is subsequently clarified in discussion), The user also has access to controls which affect the extent to which actions of mechanisms (including graphics) are linked, and affect the amount of operator participation required during the course of performing a task. Independent computational processes within the IVC central hub support operator command interaction, send commands to the system devices, and handle asynchronous feedback from system devices. The following is a brief outline of the functional capabilities implemented in the IVC architecture proper. We emphasize the IVC knowledge structure in this description; see the *Appendix* for an outline of the IVC hardware and software architectures.

System Devices: The four main system devices in the current IVC implementation are: 1) a robot controller, 2) a camera controller, 3) a video switcher, and 4) a graphics engine with geometric database. All interfaces with these system devices are in ASCII text, making it possible for the user to intervene at the lowest device level at any time if necessary. Histories of device interactions are maintained.

System Knowledge: Task-space knowledge is object-oriented, with objects distributed as appropriate *among* the system devices. Geometric knowledge is contained within the graphics database. Knowledge of mechanism kinematics resides both in the graphics system and in the appropriate controller. Coordination and conversion knowledge resides within the central control hub.

Control Actions: Actions cause motion in the task space (or simulations in the graphics space). Simple actions involve single objects within the distributed knowledge base. Multi-object and multi -manipulator semantically y primitive actions are built upon simple actions. These semantic actions, together with tasks, involving sequences of actions, are contained wholly within the central task control hub. Controls, set by the operator, affect the behavior of semantic actions and tasks. These controls turn on-and-off the linking between mechanisms associated within semantic actions and specify availability y, or Constraints on the utilization of mechanisms for automated intelligent operator assistance.

Operator Interface: The IVC user interface, **Figure 7**, provides the user with menu and interactive command line access to each of the devices, actions, semantic actions, and tasks that are defined in the system. Graphics images show available camera views at the remote site. Live video of the currently selected camera view is captured and displayed. When desired, the operator can also command and display

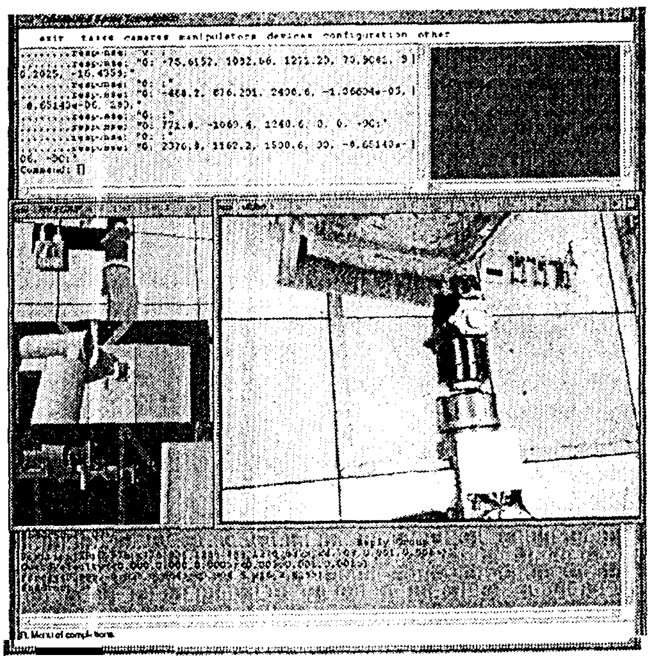

Fig 7: IVC Integrated Operator Interface

calibrated graphic views overlaid on the selected video window -- enabling a registered 3-D spatial/temporal VR-based presentation of partially obscured objects. analog task cues (plausible motion trajectories), discrete markers (control state transitions), and iconic objects (executable procedures, object information. error branches. etc.).

3.2 IVC **Knowledge Components**

Intelligent Viewing Control is based upon a semantic linking between manipulations being performed and constraints on viewing the action. For example, in the related development by Wakita et at. of MITI-Electrotechnical Laboratory on "automatic camera work [66, 67]" for pick-and-place operations in the MEISTER system [48], cameras pan and tilt to continuously view a fixed point between the fingers of the gripper. Cameras zoom-in during moves to close proximity of contact (move_to_grip) and during moves making contact (grip and put_it_on); cameras zoom-out during moves leaving close proximity (depart) or during long free-space motions (move_to_approach for place). Semantic linkage is embodied within object methods implementing pick-and-place. By contrast, in Intelligent Viewing Control, the semantic linkage between manipulation actions and corresponding camera actions is *implemented within two separate parts of/he control architecture.* This architectural (and syntactic) decoupling of actions enables the generalization of camera control from scripted object-manipulator viewing behaviors to more general cases where *camera actions at a given point in the task sequence can be made context-dependent [38, 47]* on prior, current, and posterior task knowledge, including the task interaction constraints (static and dynamic geometric contact, viewing obscuration, etc.).

Task dependent "fixtures [49]," viz. sets of one or more reference frames associated with an object to be acted upon, are implemented within the geometric database portion of the object. These fixtures provide the knowledge needed by the intelligent behaviors within the remote robot controller. They provide information about the location and spatial scope of the behavior, thereby specifying constraints on the focus of attention appropriate for viewing the behavior. Semantic actions, which link objects with actions of mechanisms and link actions of multiple mechanisms, contain procedural information for intelligent viewing, Each *semantic action is* a multi-step procedure involving graphical, camera, and robotic devices:

- choose focus of attention
- simulate action in graphics
- command camera motion
- command robot motion
- update graphics per robot motion feedback

Camera motions, per activities of the robot workcell shown **Figure 8** upper right (and Figure 1), may either be completed **before.** or commanded concurrent with robot motion. The focus of attention may be the robot manipulator, a tool, or an object being worked on. Steps within the sequence may be enabled or inhibited via controls set by the user.

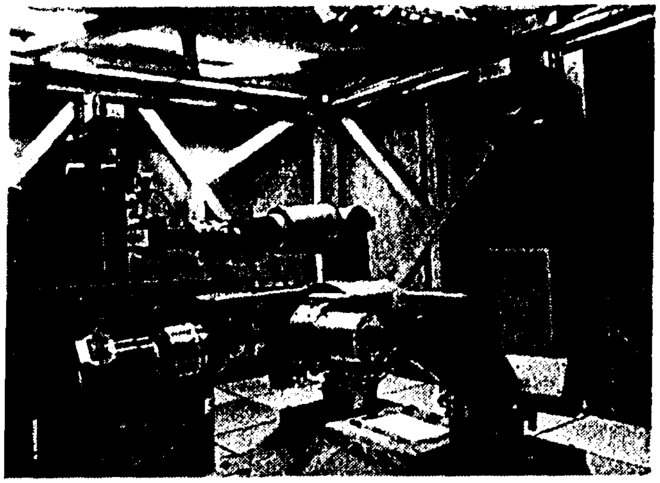

Fig 8: Multi-Camera Workcell of JPL Telerobotic System (cameras overhead, back wall, side view, and wrist)

4 Some Open Issues

The stated theme of this workshop is development of intelligent control for large, complex systems -- ones that are intrinsically difficult to analyze, and predictively synthesize, as Simon comments in *[58].* Developmental issues, per the workshop call, include "drawbacks in broad use of arbitrary heuristics, declaration of voluntary rules and principles, lack of integration between discrete mathematics and continuous control, lack of a unifying theoretical framework, *etc.* " Telerobotic systems seem to qualify well for these considerations. Telerobots, in idealization, utilize the best of human and robot skills to expand human capability and extend its presence to remote sites, or those not conventionally accessible due to state or environment. A true human-machine telerobotic design synergy would draw on the hierarchically structured cognitive, perceptual, and motor skills [46] of each agent, as appropriate to the task. Further, such distributed system designs could invoke cooperative multiple agents, human or robotic. References [15, 37] give a sense of conceptual dichotomies that arise in attacking such issues.

The discussions and case studies we presented in previous sections are illustrative, indicating the scope of telerobotics, and the numerous operational challenges that confront it. Many issues remain open for the further development of robotics, human-machine interfaces, and cognitive and computational constructs to unify these first two areas. Looking beyond the intrinsically interesting problems of robot mechanisms and their dexterous controls, we highlight a few fundamental research issues:

Uncertainty Management: all robotic tasks are inherently based on knowledge that is partial, priced. and errorful. This problem can be attacked on several fronts. one being the explicit modeling of perceptual and control uncertainties and associated recovery strategies -- e.g., geometric, probabilistic. or other representational constructs and error correcting plans/controls [51 -- also emulation of certain biophysical and cybernetic principles in simpler life forms that capture useful reflexive and sensor-motor behaviors [23. 37]. Whatever the approach, failure to explicitly account for spatial and temporal variabilities in task execution lead to systems that arc inherently brittle, as well as unpredictable.

Task Representation: Modeling for a large class of systemic prototypes and procedures is computationally impractical in large domains -- and as some have argued (e.g., see the edited collection of Maes op. cit. [14]) -- fundamental] y inconsistent with biological evidence and efficacy. In a crude sense, this issue is a deterministic variation on the one above: either to model in categorical detail for task structure and actions, or to model distributionally for priors and ensemble outcomes. In that approximate context, *learning* strategies [2] become explicitly important. While learning approaches can finesse certain implicit representational or associative behavioral issues, such approaches still confront underlying problems of search space and training set dimensionality in tasks of rich content.

Heuristics, Linguistics & Unification: robotic automation and telerobotic operations are grounded in a physical domain, one often casually structured. Defining robust, well-modeled approaches to even factory floor automation problems [5, 45] is difficult, and tends to promote "cooperative/ reduced complexity" task design, to good economic effect. Analyzing and synthesizing (tele)robotic *system-level* task performance has to date been a largely empirical undertaking, and points strongly to the importance of experimentation in robotics. There are few "clean" problems in the sense of AI system structure, bounded reasoning, or hierarchical and multi-resolution design practices [46, 65]. Along with the two issues above, this is certainly one reason to turn the tables, utilizing the robot sensor and effecter interfaces and task environment itself to build behaviorally driven, subsumptive, and ultimately emergent systems [15, 37]. Still, it is in the nature of scientific inquiry and engineering design to seek predictive, explanatory models, 10 do will require tools of analysis and synthesis that can make explicit the invariant properties of natural task structures, their perceptual features, the universal transformations (syntax) acting on objects and agents, and an unambiguous task semantics

Acknowledgments

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Appendix: IVC Implementation

A.1 Telerobot Hardware Configuration [12, 43, 50]

Local Site (Operator): At the "local" site (pictured in Figure 1), the operator is offered 3 color video monitors and 2 workstations (by Silicon Graphics and Sun Microsystems) for viewing. The graphics workstation is equipped with VideoLab in order to display live video on its monitor. The Sun Spare 10 contains an XVidco™ board by parallax and provides the capabilities to display live video on the Sun monitor and to perform hardware-based JPEG compression of images. Video signals are routed to the monitors and workstations through a computer-controlled 8x8 video switch. In addition to the conventional computer interfaces (GUIs, mouse, voice, etc.), the operator has at his disposal a SpaceBall and dual 6 degree-of-freedom (dof) Force Reflecting Hand Controllers. The FRHC's arc controlled at the high level by a *single* VME system (running the VxWorks Real-time OS) which is memory -mapped to servo level Universal Motor Controllers (uMC).

Remote Site (Robot): The "remote" site (pictured in Figure 8) consists of two comprehensive manipulation and viewing systems. Dual redundant 8-dof AAI robots are each equipped with a force sensing, parallel jaw instrumented "Smarthand" (JPL design with integrated force-torque F/T sensor, on-board signal pre-processing, etc.). A VME system (running VxWorks) controls these robots in task space and is memory -mapped to the joint-level servo controllers (UMC's). The remote VME system also controls the viewing gantry, which consists of three positionable camera platforms (4-dof each: 2-transl at ion axes, and pan-tilt) with four cameras (two being a stereo pair) having computer controllable focus, iris, and zoom. A fifth camera is optionally mounted on one Smarthand for "eye-in-hand" viewing.

A.2 IVC **Software Environment**

Figure A1, next page, summarizes the physical organization of the telerobot system with which IVC interacts. The IVC central hub is implemented in Allegro Common Lisp with the Common Lisp Interface Manager, Version 2 (CLIM-2). The Deneb Robotics, lnc., TELEGRIP™ package is used for the graphics simulation and geometric database support. The video switch interface, camera controller, and robot controller are all implemented in custom C software.

The central hub is connected to the four system devices via sockets, with ASCII text interfaces to each. The TELEGRIP interface utilizes the pre-defined CL1 command language and protocol: each command to TELEGRIP is followed by a single response. The video switcher interface also has an alternating command/response protocol. The camera controller interface is currently driven open loop, with commands to the controller, but no responses. The robot controller interface is asynchronous: Commands are sent in groups, with execution of motion commands starring upon the receipt of a numbered EndGroup command. During robot motion, a stream of position and force feedback data are returned. An independent process within the central hub continuously monitors robot controller feedback, maintains the latest robot force and position states and monitors for command group completion status. This process also updates the graphics display. if TELEGRIP is connected and graphics update upon feedback control is enabled.

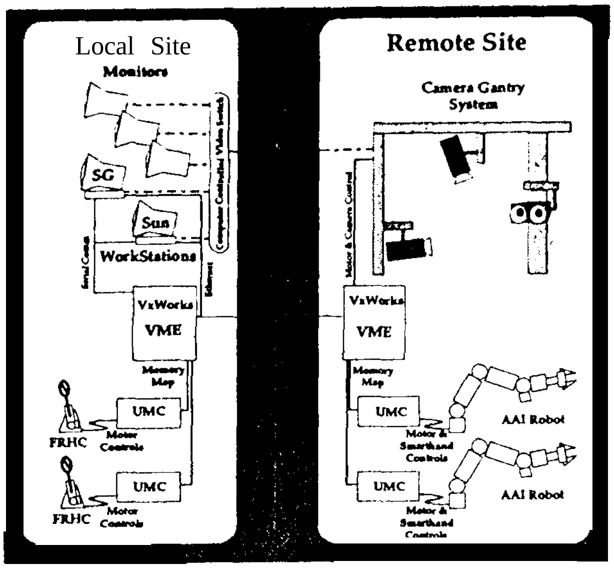

Fig. A 1: Physical Architecture of IVC Environment

TELEGRIP models had previously been developed for most of the JPL robotics lab workcell objects and mechanisms; new *models* were created for the 2-DOF translational camera positioning mechanisms. The pan/tilt heads were not modelled; the camera is modelled as a point at the intersection of the pan and tilt actuator axes, Manipulation fixtures are modelled as tags attached to paths, which in turn are attached to parts and devices in the TELEGRIP terminology.

The supporting central hub software consists of 1) the CLIM menu interface, 2) connections to system devices, 3) extensions to the robot task object models, 4) semantic action functions, and 5) system control parameters. The CLIM interface provides mouse and keyboard command access to CLIM command functions. These functions can be called by any Common Lisp functions, including other CLIM commands. This permits the layering of multiple levels of access and intelligence within the same interface. Device models consist of the active socket object plus device-specific control parameters. One such parameter enables or inhibits the automatic updating of graphics from robot controller feedback data. Device models also maintain command and response histories,

"Simple actions" are functions which generate appropriate teat suing commands 10 be sent across the socket interface to the device. "Semantic actions" are functions which assemble the appropriate data and call the appropriate set of simple actions. Semantic action functions follow the CLIM command function structure, making them available to the operator through the menu and command interface. Semantic actions communicate all information needed by the device to perform its portion of the task. For example, camera and robot move commands within the TELEGRIP CLI language reference a previously selected "current" camera and "cur-mot" *robot* which are internal TELEGRIP state information. Relevant internal states are actively set by each semantic command to ensure its executional integrity, This conservative, semanticall y complete approach to communications is necessary to implementation of a system which allows operator intervention at multiple levels of abstraction.

The cameras are the only objects which have a significant portion of their data maintained within the central hub. The models contain the current camera translational positions and the pan & tilt angles. The camera objects also compute the pan and tilt angles required for viewing an object of interest. Other information such as translational range of motion and the name of the camera in each of the TELEGRIP and camera controller interfaces is also maintained.

All communication of position information throughout the system is in millimeters x, y, and z relative to a global world frame. Communication of orientation information is in degrees of *roll*, pitch, and yaw with respect to TELEGRIP and in radians with respect to the robot controller. Camera control commands are given in millimeters of translation in device prismatic joint space, and in radians of pan and t ilt.

References

1) *Tutorial on Robotics* (Eds., C.S.G. Lee, R.C. Gonzales, and K .S. Fu). IEEE Computer Society Press, 1986 (2nd edition): Washington, D.C.

2) *Toward Learning Robots* (Ed., W. Van de Welde). The MIT Press, 1993: Cambridge, MA.

3) *Virtual Reality: Scientific and Technological Challenges,* a report of the Committee on Virtual Reality Research & Development (Chr., N. Durlach), **National Research** Council. NAS Press, 1994: Washington D.C.

4) *Space Station/OACT Robotics Technology Study,* " Vol. 1. (ref: NASA TM 11931}SS-039, Oceaneering Space Systems with McDonnell Douglas Aerospace, 1993).

5) *Geometric Uncertainly in Motion Planning: Summary Report and Bibliography,* an NSF sponsored workshop, report issued as IRIS TR No, 297, University of Southern California, Los Angeles, CA (Eds., K. Y. Goldberg, M. T. Mason, and A. Requicha), August, 1992.

6) S. Abbrams and P. K. Allen, "Sensor planning in an active robotic workcell," Proc. DARPA Image Understanding Workshop. January, 1992 (Morgan-Kaufmann).

7) J. S. Albus, *Brains, Behavior, and Robotics.* McGraw-Hill, 1981: New York, NY.

8) J. S. Albus, R. Lumia, and H. McCain, "Hierarchical control of intelligent machines applied to Space Station robots." IEEE Trans. Aerospace and Electronic Systems. Vol. 24, No. 5. pp. 535-541, September, 1988.

9) J. S. Albus and R. Quintero, "Toward a reference model architecture for real-lime intelligent control systems (ARTICS)," in Robotics and Manufacturing, Vol. 3 (Proc. ISRAM '90). ASME Press, 1990: New York, NY.

10) P. G. Backes and K. S. Tso, "UMI: an interactive supervisory and shared control system for telerobotics," in Proc. 1990 IEEE Intl. Conf. Robotics and Automation, Cincinnati, OH, May.

11) P. Backes, J. Beahan, and B. Bon, "Interactive command building and sequencing for supervised autonomy," in Proc. 1993 IEEE IntJ. Conf. Robotics and Automation, Atlanta May.

12) A. K. Bejczy and Z. F. Szakaly, "An 8 d.o.f. dual arm system for advanced teleoperation performance experiments," in Proc. SOAR '91 Symposium (Space Operations, Applications, and Research), Houston, TX, July 1991.

13) A. K. Bejczy, P. Fiorini, W. S. Kim, and P. Schenker, "An integrated operator interface for advanced teleoperation with time-delay," in Proc. 1994 IEEE/RSJ/ Cd Intl. Conf. IROS, Munich, Germany, September, 1994.

14) R. A. Brooks, "Elephants don't play chess," pp. 3-16, in *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back* (Ed., P. Maes). The MIT Press, 1990: Cambridge, MA.

15) R. A. Brooks, "Building brains for bodies," in Autonomous Robots, Vol. 1, no. 1, pp. 7-26, November, 1994.

16) IL A. Brooks, "Intelligence without representation," and P. S. Rosenbloom, J. E. Laird, A. Newell, and R. McCarl, "A preliminary analysis of the Soar architecture as basis for general intelligence," in *The Foundations of Artificial Intelligence* (Ed., D. Kirsch). The MIT Press, 1992: Cambridge, MA.

17) T. Brooks, I. Ince, and G. Lee, "Vision issues for space teleoperation assembly and servicing (VISTAS)," Rept. No. STX/ROB/91 -01, (NASA Contract #NAS-5-30440), STX Corporation, Lanham, MD, January, 1991.

18) B. K. Christensen, W.R. Drotning, and S. Thornborg, "Graphical model based control of intelligent robot syswins,"' Proc. 1991 IEEE Intl. Conf. Systems. Man, and Cybernetics, Sacramento, CA. May.

19) L. Conway, R. Volz, and M. Walker, "Teleautonomous systems: methods and architectures for intermingling autonomous and telerobotic technology," in Proc. 1987 IEEE Intl. Conf. Robotics and Automation, Raleigh, NC, March.

20) H. Das, P. S. Schenker, H. Zak and A. K. Bejczy, "Teleoperated satellite repair experiments," in Proc. 1992 IEEE/RSJ Intl. Conf. IROS, Raleigh-Durham, NC, July.

21) H. Das,, H. Zak, W. S. Kim, A. K. Bejczy, and P. S. Schenker, "Operator performance with alternative manual modes of control," Presence, vol. 1, no. 2, pp. 201-218, Spring 1992.

22) P. Fiorini, A. K. Bejczy, and P. S. Schenker, "Integrated interface for advanced teleoperation," IEEE Control Systems, Vol. 13, no. 5, pp. 15-20, October, 1993,

23) A.M. Flynn, R. A. Brooks, W.M. Wells, 111, D. S. Barrett, "Intelligence for miniature robots," Sensors and Actuators, Vol. 20, pp. 187-196, 1989.

24) J. Funda, T. S. Lindsay, and R. P. Paul, "Teleprogramming: toward delay-invariant remote manipulation," Presence, Vol. 1, no. 1, pp. 29-44,192.

25) G. Harbaugh, "An operational perspective on the use of telerobotic systems in manned spaceflight," in Proc. AIAA Conf. Space Programs and Technologies, Paper 92-1450, Huntsville, AL., March, 1992.

26) S. Hirai, T. Sato and T. Matsui, "Intelligent and cooperative control of telerobot tasks," IEEE Control Systems Magazine, pp. 51-56, February, 1992.

27) G. Hirzinger, B. Brunner, J. Dietrich, and J. Hendl, "Sensor-based space robotics -- ROTEX and its telerobotic features," IEEE Trans. Robotics and Autom., Vol. 9, No. 5, pp. 649-663, October 1993.

28) G. Hirzinger, J, Heindl, K. Landzettel, and B. Brunner, 'Multisensory shared autonomy -- a key issue in the space robot technology experiment ROTEX," Proc. 1993 IEEE/RSJ Intl. Conf. IROS, Raleigh-Durham, NC, July.

29) W. S. Kim and P. S. Schenker, "Teleoperation training simulator with visual and kinesthetic force reality," in Human Vision, Visual processing, and Visualization, Proc. SPIE 1666, San Jose, CA, February, 1992.

30) W. S. Kim, P. S. Schenker, A. K. Bejczy, S. I-cake, and S. Ollendorf, "An advanced operator interface design with preview/predictive displays for ground-controlled space telerobotic servicing," in Telemanipulator Technology and Space Robotics, Proc. SPIE 2057, Boston, MA, September, 1993

31) W. S. **Kim.** P. S. Schenker, A. K. Bejczy and S. Hayati, "Advanced graphic interfaces for telerobotic servicing and inspection.'" in Proc. **1993 IEEE-RSJ Intl. Conf. IROS,** Yokohama. Japan. July.

32) W. S. **Kim,** "Virtual reality calibration for telerobotic servicing," in Proc. 1994 IEEE Intl. Conf. Robotics and Automation. San Diego. CA, May.

33) P. Lee. A. K. Bejczy, P. S. Schenker, and B. Hannaford, "Telerobot configuration editor," in Proc. IEEE Intl. Conf. Systems, Man, and Cybernetics, Los Angeles. CA, November, **1990.**

34) **S. Lee,** P.S. Schenker, and J. Park, "Sensor-knowledge-command fusion paradigm for man/machine systems," Proc. IFAC Distributed Intelligence Systems '91, Arlington, VA, August, 1991.

35) M. J. **MacDonald** and R. D. Palmquist, "Graphical programming: on-line robot simulation for telerobotic control," in Proc. Intl. Robots and Vision Automation Conf., Detroit, MI, April, 1993.

36) **M. Mallem,** F. Chavand, and E, Cone, "Computer-assisted visual perception in teleoperated robotics," Robotics Vol. 10. PP. 93-103.1992.

37) **D.** McFarland and T. Bosser, *Intelligent Behavior in Animals and Robots.* The MIT Press, 1993: Cambridge, MA.

38) G. **T.** McKee and P. S. Schenker, "Visual acts for goal-directed vision," in **1995** IEEE Intl. Conf. Robotics and Automation, Nagoya, Japan, May.

39) G. **T.** McKee and P. S. Schenker, "Human-robot cooperation for automated viewing during teleoperation," in *1995* IEEE-RSJ Intl. Conf. IROS, Pittsburgh, PA, August.

40) G. **T.** McKee and P. S. Schenker, "Networked robotics, cooperating agents, and an application to 'Visual Acts'," in Proc. IEEE/Intl. Symp. Intell. Control Workshop (Architectures for Semiotic Modeling and Situation Analysis in Large Complex Systems, Orgs.: J. Albus, A. Meystel, D. Pospelov, T. Reader), Monterey, CA, August, 1995,

41) **K. Mochizuki, 11.** Takemura, and F. Kishino, "Object manipulation and layout in 3-D virtual space using a combination of natural language and pointing," in Sensor Fusion V, Proc. SPIE 1828, Boston, MA, November, 1992.

42) **E.** Oyama, N. Tsunemoto, S. Tachi, and Y. Inoue, "Experimental study on remote manipulation using virtual reality," Presence, Vol. 2, No. 2, pp. 112-124, Spring, 1993.

43) E. **D. Paljug** and P. S. Schenker, "Advanced Teleoperation Control Architecture," in Telemanipulator Technology and Space Robotics, Proc. SPIE 2057. Boston. MA, **September, 1993.**

44) **R. P. Paul, C.** Sayers, and M. R. Stein, "The theory of teleprogramming," Jrnl. Robotics SoC. Japan, Vol. 11, No. 6.1993.

45) **R. J. Popplestone,** Y. Liu, and R. Weiss, "A group theoretic approach to assembly planning." AI Magazine. pp. 82-97, Spring 1990.

46)]. **Rasmussen, "Skills. rules,** and knowledge: signals, signs, and symbols, and other distinctions in human performance models," IEEE Trans. Systems, Man, and Cybernetics, SMC-13, pp. 257-266, 1983.

47) R. **D. Rimey** and C. D. Brown, "Sequences. structure. and active vision," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, Lahaina, Maui (HI), 1991.

48) T. Sato and S. Hirai, "MEISTER: A model enhanced and skillful teleoperational robot system," in Proc. 4th Intl. Symp. Robotics Rsch. (R. Bolles and B. Roth, Eds.), 1987.

49) C. Sayers and R. Paul, "Synthetic fixturing," in Proc. Haptic Interfaces for Virtual Environments and Teleoperator Systems, ASME Winter Annual Meeting, New Orleans, LA, November, 1993.

50) **P.** S. **Schenker,** A. K. Bejczy, W. S. Kim, and S. Lee, "Advanced man-machine interfaces and control architecture for dexterous teleoperations," in Proc. Oceans '91, pp. 1500-1525, Honolulu, HI, October, 1991.

51) **P.** S. **Schenker,** "Intelligent robots for space applications," pp. 545-591 in *Intelligent Robotic Systems: Analysis, Design, and Programming (S. Tzafestas, Ed.).* Marcel Dekker, 1991: New York, NY.

52) P. S. **Schenker** and W. S. Kim, "Remote robotic operations and graphics-based operator interfaces," in Robotics and Manufacturing, Vol. 5 (Proc. ISRAM '94). ASME Press, 1994: New York, NY.

53) 1'. S. **Schenker, S. F.** Peters, E. D. Paljug, and W.S. Kim, "Intelligent viewing control for robotic and automation systems," in Sensor Fusion VII, Proc. SPIE 2355, Boston, MA, October, 1994.

54) **P. S. Schenker** and S. Hirai, "A US-Japan collaborative robotics research program," in Proc. Third Intl. Symp. on AI, Robotics, and Automation for Space (i-SAIRAS'94), Pasadena, CA, October, 1994.

55) **B. Schneiderman,** *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* Reading, MA: Addison-Wesley, 1987.

56) 1'. **Sheridan,** *Telerobotics, Automation, and Human Supervisory Control.* The MIT Press, 1992: Cambridge, MA.

s7) T. Sheridan, "Space teleoperation through time delay: review and prognosis." IEEE Trans. Robotics & Autom., Vol. 9. No. 5. pp. 592-606, October. 1993.

58) H. A. Shots. "Prediction and prescription in systems modeling, " operations Research. Vol. 38. No. 1, pp. 7-14. Jan-Feb, 1990.

59) D. Simon, B. Espiau, and K. Kapellos, "Computcr-aided design of a generic robot controller handling reactivity and real-time control issues." IEEE Trans. Control Sys. Techn., Vol. 1. No. 4, December. 1993.

60) M.R. Stein and R.P. Paul. "Operator interface for time-delayed teleoperation, with a behavior-based controller." Proc. IEEE Intl. Conf. Robotics & Automation, San Diego, CA, May, 1994.

61) Matthew R. Stein, Behavior-Based Control for Time-Delayed Teleoperation, Ph.D. Thesis (April, 1994), The University of Pennsylvania, Philadelphia (Advisor: R.P. Paul; NASA Technical Advisor: P. S. Schenker, Jet Propulsion Laboratory).

62) M. Stein, R. Paul, E. Paljug, and P. Schenker, "A cross-country teleprogrammi ng experiment." in 1995 IEEE-RSJ Intl. Conf. IROS, Pittsburgh, PA, August.

63) T. Takahashi and H. Ogata, "Robotic assembly operations based on task-level teaching in virtual reality," in Proc. 1992 IEEE Intl. Conf. Robotics and Automation, Nice, France, May.

64) K. A. Tarabanis, p. K. Allen, and R. Y. Tsai, "A survey of sensor planning in computer vision," IEEE Trans. Robotics and Autom., Vol. 11, no. 1, pp. 84-104, February, 1995.

65) K. J. Vicente and J. Rasmussen, "Ecological interface design: theoretical foundations," FEEE Trans. Systems. Man, and Cybernetics, SMC-22, pp. 589-606, 1992.

66) Y. Wakita, S. Hirai and T. Kino, "Automatic camera-work control for intelligent monitoring of telerobotic tasks," in Proc. 1992 IEEE/RSJ Intl. Conf. IROS, R{sleigh-hrrham, NC, July.

67) Y. Wakita and S. Hirai, "Hierarchical control of a visual monitoring system for telerobot tasks," in Proc. 1993 IEEE/RSJ Intl. Conf. IROS, Yokohama, Japan, July.

68) Y. Wakita, S. Hirai, K. Machida, S. Peters, E. Paljug, and P. Schenker, "Telerobotic task execution over Internet," submitted to 1995 IEEE-RSJ Intl. Conf. IROS, Pittsburgh, PA, August.